

# SPIN GLASS COMPUTATIONS AND RUELLE'S PROBABILITY CASCADES

LOUIS-PIERRE ARGUIN

**ABSTRACT.** We study the Parisi functional, appearing in the Parisi formula for the pressure of the SK model, as a functional on Ruelle's Probability Cascades (RPC). Computation techniques for the RPC formulation of the functional are developed. They are used to derive continuity and monotonicity properties of the functional retrieving a theorem of Guerra. We also detail the connection between the Aizenman-Sims-Starr variational principle and the Parisi formula. As a final application of the techniques, we rederive the Almeida-Thouless line in the spirit of Toninelli but relying on the RPC structure.

## 1. INTRODUCTION

The Sherrington-Kirkpatrick (SK) model is a mean-field spin glass system on configurations  $\sigma \in \{\pm 1\}^N$  of  $N$  spins with the Hamiltonian

$$H_N(\sigma) = \frac{-1}{\sqrt{N}} \sum_{1 \leq i < j \leq N} J_{ij} \sigma_i \sigma_j + h \sum_{1 \leq i \leq N} \sigma_i.$$

The couplings  $J_{ij}$  are independent standard gaussian variables and  $h \in \mathbb{R}$ .

It is now a theorem that the quenched pressure of the SK model in the thermodynamic limit,  $P_{SK}(\beta, h) := \lim_{N \rightarrow \infty} \frac{1}{N} E_J [\log \sum_{\sigma} e^{-\beta H_N(\sigma)}]$ , is given by the celebrated Parisi formula [13]:

$$P_{SK}(\beta, h) = \inf_{x(\cdot)} \left\{ \log 2 + f_x(0, h) - \frac{\beta^2}{2} \int_0^1 qx(q) dq \right\}.$$

The infimum is over all increasing, right-continuous functions  $x : q \mapsto x(q)$  on  $[0, 1]$  such that  $x(0) = 0$  and  $x(1) = 1$ . At the heart of this formula is the so-called *Parisi functional*  $x \mapsto f_x(0, h)$  where  $f_x(q, y)$  is the solution to the partial differential equation

$$(1) \quad \partial_q f(q, y) + \frac{1}{2} \left[ \partial_y^2 f(q, y) + x(q) (\partial_y f(q, y))^2 \right] = 0$$

with the boundary condition  $f(1, y) = \log \cosh(\beta y)$  [10, 14].

A different approach in computing the pressure of the SK model was taken by Aizenman, Sims and Starr (*AS<sup>2</sup>*) [2]. In this approach, the pressure is expressed through a general variational principle over *random overlap structures* (ROSt). A ROSt is a measure  $\mu$  on a pair  $(\xi, Q)$  where  $\xi = \{\xi_{\alpha}\}_{\alpha \in \mathcal{A}}$  is a set of weights labeled by  $\mathcal{A}$  and  $Q = \{q_{\alpha, \alpha'}\}$  is a positive semi-definite form on  $\mathcal{A}$ . The variational principle uses functionals on ROSt's of the form

$$(2) \quad E_{\mu} \left[ \log \frac{\sum_{\alpha} \xi_{\alpha} e^{\psi(\eta_{\alpha})}}{\sum_{\alpha} \xi_{\alpha}} \right]$$

where  $\eta$  is a gaussian field with covariance  $Q$  and  $\psi$  is a specific function.

*Date:* June 12th, 2006.

Work supported in part by NSERC and FQRNT postgraduate fellowships and NSF grant DMS 0602360.

As it was pointed out in [2, 3], the link between the two approaches in computing the SK pressure is provided by a particular family of ROST's known as the *Ruelle's Probability Cascades* (RPC). In fact, the Parisi formula is retrieved by restricting the  $AS^2$  variational problem to this class of ROST's. These ROST's possess two important features. First, their overlap matrix can be represented as a tree structure, sometimes qualified as *ultrametric*. Second, these processes are stable under stochastic shift of a certain kind. This property shall be defined precisely and be referred to as the *quasi-stationarity property*.

The goal of this paper is to exhibit specific techniques for spin glass computations with the RPC's. The RPC is a natural setting to study the Parisi functional and its properties and to perform computations relevant to the SK model.

The paper is organized as follows. We first introduce the RPC. We then precisely define the quasi-stationarity property and give sufficient and necessary conditions on the stochastic shift for stability. We proceed by studying functionals on RPC's of the form (2) that reduce to Parisi-like functionals. We derive differentiation formulas for these functionals which naturally lead to continuity and monotonicity properties thereby retrieving a theorem of Guerra [6] for the Parisi functional. Finally, we detail the connection between the  $AS^2$  variational principle restricted to the class of RPC's and the Parisi formula for the pressure of the SK model. As an example of computation for this variational principle, we prove the instability of the high-temperature solution above the Almeida-Thouless line following the idea of Toninelli [15] but using the properties of the RPC functionals introduced earlier. It must be emphasized that the key element involved in most calculations is the quasi-stationarity property of the RPC.

The connection between the RPC's and the SK model is still to be fully understood. In particular, the question of whether or not the overlap distribution of the SK model is supported on ultrametric matrices is open. The quasi-stationarity property of the RPC seems to be at the core of this question as pointed out in [2, 3, 7].

The author is thankful to Michael Aizenman for introducing him to the subject and for plenty of insightful discussions.

## 2. RUELLE'S PROBABILITY CASCADES

*Ruelle's Probability Cascades*, or RPC's, are cascades of Poisson point processes which carry a natural hierarchal distance between the atoms of the cascade. The RPC was formulated by Ruelle based on the *Generalized Random Energy Model* (GREM) originally defined by Derrida as a limit of finite point processes [5]. Ruelle's formulation extends GREM's to allow continuous branchings or hierarchies [11]. In these notes, we are interested in the case of finite number of branching levels where the definitions of GREM and RPC correspond. Keeping this in mind, we will often use the word GREM for an RPC with a finite number of branching levels.

**2.1. Probability Measures on  $[0, 1]$ .** We start by fixing the notation that will be needed in the definition of the GREM and used throughout the paper.

Let  $\mathcal{M}$  be the set of probability measures on  $[0, 1]$  and  $\mathcal{M}_a \subset \mathcal{M}$ , the subset of atomic measures with finite number of atoms. For later purposes, we also introduce  $\mathcal{M}_{a,k}^{<1}$ , the subspace of  $\mathcal{M}_a$  with exactly  $k+1 < \infty$  atoms, one of them located at 1 and the remaining located on  $[0, 1)$ . We write  $\mathcal{M}_a^{<1}$  for  $\bigcup_{k \in \mathbb{N}} \mathcal{M}_{a,k}^{<1}$ .

The space  $\mathcal{M}_{a,k}^{<1}$  corresponds to the following subspace of  $[0, 1]^{k+1} \times [0, 1]^{k+1}$  most commonly used in the spin glass literature. We associate to  $x \in \mathcal{M}_{a,k}^{<1}$  the pair  $(\mathbf{x}, \mathbf{q})$  where

$\mathbf{x} = (x_i, i = 1, \dots, k+1)$  and  $\mathbf{q} = (q_i, i = 1, \dots, k+1)$  with the constraints

$$\begin{aligned} 0 &< x_1 < x_2 < \dots < x_{k+1} \equiv 1 \\ 0 &\leq q_1 < q_2 < \dots < q_{k+1} \equiv 1. \end{aligned}$$

We also set  $q_0 = 0$ . In this notation,  $q_i$  refers to the position of the  $i$ -th atom and  $x_i = x(q_i)$ .

Throughout these notes, we will identify a probability measure with its distribution function and write  $x \in \mathcal{M}$  for a distribution function  $x$  of a measure in  $\mathcal{M}$ . We will be naturally led to endow  $\mathcal{M}$  with the topology induced by the  $L^1([0, 1], g'(q)dq)$ -norm on the distribution functions for a given smooth function  $g$ . It turns out that the topology does not actually depend on  $g$ . We refer to this topology as the  $L^1$ -topology on  $\mathcal{M}$ . In fact, the  $L^1$ -topology is simply the weak topology on  $\mathcal{M}$  (see Appendix A for details). Note that the subset  $\mathcal{M}_a^{<1}$  is dense in  $\mathcal{M}$  in the  $L^1$ -topology.

**2.2. The REM.** The building block of the cascade of the GREM is the simple REM point process.

**Definition (REM).** Let  $0 < x < 1$ . A  $REM(x)$  is a Poisson point process on  $\mathbb{R}^+$  with intensity measure  $xs^{x-1}ds$ .

The value of the parameter  $x = 1$  is evidently singular. It will become clear later that this is the fundamental reason for introducing the subspace  $\mathcal{M}_a^{<1} \subset \mathcal{M}_a$ . Fortunately, we will later consider functionals of the point process that allow a continuous extension to the case  $x = 1$ .

Let  $\zeta = \{\zeta_\alpha\}$  be a  $REM(x)$ . It is not too hard to show that  $\sum_\alpha \zeta_\alpha < \infty$  a.s. (see e.g. [11]). In particular,  $\zeta$  is locally finite on  $(0, \infty)$  and bounded on the right almost surely. Therefore, it is possible to enumerate the points of a realization in decreasing order i.e.  $\zeta_1 > \zeta_2 > \dots$

It turns out that the REM possesses an interesting stability property under stochastic shift. This property is at the root of the techniques and results presented in this paper. Let  $\zeta$  be a  $REM(x)$ . We consider a random variable  $W$  on  $\mathbb{R}^+$  with distribution  $\nu$  such that  $E_\nu[W^x] = \int_0^\infty w^x d\nu(w) < \infty$ . A proof of the following can be found in Proposition 3.1 of [12].

**Proposition 1** (Quasi-Stationarity of the REM). Let  $\zeta = \{\zeta_i\}_{i \in \mathbb{N}}$  be a  $REM(x)$  and  $W$  be as above. Consider  $\{W_i\}_{i \in \mathbb{N}}$  iid  $W$ -distributed and independent of  $\zeta$ . Define the point process  $\tilde{\zeta} := \{\zeta_i W_i\}_{i \in \mathbb{N}}$ . The following hold

(1) Quasi-Stationarity:  $\tilde{\zeta}$  is a  $REM(x)$  scaled by  $E_\nu[W^x]^{1/x}$ , i.e.

$$\tilde{\zeta} \stackrel{\mathcal{D}}{=} E_\nu[W^x]^{1/x} \zeta.$$

(2) Backward Shift: Let  $\{\tilde{\zeta}_j\}$  be ordered in decreasing order. Let  $\pi : \mathbb{N} \rightarrow \mathbb{N}$  be the random permutation induced by the random shift, i.e.  $\pi(i) = j$  iff  $\tilde{\zeta}_j = \zeta_i W_i$ . Then,  $\{W_{\pi^{-1}(j)}\}_{j \in \mathbb{N}}$  are iid and independent of  $\tilde{\zeta}$  with distribution

$$\frac{w^x d\nu(w)}{E_\nu[W^x]^{1/x}}.$$

**2.3. The GREM.** We now construct the GREM process as a cascade or hierarchy of REM's. We start by defining the point process associated to a GREM, we then introduce the overlap matrix induced by the cascade. We choose to identify the class of GREM processes with the space of atomic measures  $\mathcal{M}_a^{<1}$ . Therefore, elements of the class of GREM's are distinguished by the choice of  $x_l$ 's but also by the choice of overlap parameters  $q_l$ 's. This is a useful identification as functionals over GREM's become functionals on a dense subspace of the space  $\mathcal{M}$  of probability measures on  $[0, 1]$ .

Let  $x \in \mathcal{M}_{a,k}^{<1}$  with atoms at  $\{q_l\}_{1 \leq l \leq k+1}$  and  $x(q_l) = x_l$ . Recall that  $q_{k+1} = 1$  and  $x_{k+1} = 1$ . Consider  $\alpha \in \mathbb{N}^k$ ,  $\alpha = (\alpha_1, \dots, \alpha_k)$ . It is convenient to define, for  $l = 0, \dots, k$ , the truncation  $\alpha(l) := (\alpha_1, \dots, \alpha_l)$ . By convention,  $\alpha(0) = 0$ . Consider for each  $l = 1, \dots, k$  a collection of independent REM( $x_l$ ) indexed by  $\alpha(l-1) \in \mathbb{N}^{l-1}$

$$\left( \zeta^{\alpha(l-1)}, \alpha(l-1) \in \mathbb{N}^{l-1} \right).$$

The notation  $(\zeta^{\alpha(l-1)})_j$  will designate the  $j$ -th point of a realization of the process  $\zeta^{\alpha(l-1)}$ . By the convention  $\alpha(0) = 0$ , there is only one process in the collection  $l = 1$ .

We define recursively a hierarchy of point processes  $\{\xi_{\alpha(l)}\}_{\alpha(l) \in \mathbb{N}^l}$  at levels  $0 \leq l \leq k$  as follows:  $\xi_{\alpha(0)} := 1$  and  $\xi_{\alpha(l)} := \xi_{\alpha(l-1)} (\zeta^{\alpha(l-1)})_{\alpha_l}$ . The resulting point process is then

$$\xi_\alpha := \xi_{\alpha(k-1)} (\zeta^{\alpha(k-1)})_{\alpha_k}.$$

To keep track on the branching information, it is useful to consider the filtration  $\mathcal{F}^\xi = (\mathcal{F}_l^\xi, 0 \leq l \leq k)$  with  $\mathcal{F}_0$  being the trivial  $\sigma$ -algebra and  $\mathcal{F}_l^\xi = \sigma(\{\xi_{\alpha(l')}\}_{\alpha(l') \in \mathbb{N}^{l'}, 0 \leq l' \leq l})$ . Here,  $\sigma(\cdot)$  designates the  $\sigma$ -algebra generated by the collection of variables therein.

The overlap matrix of the cascade  $Q = \{q_{\alpha, \alpha'}\}_{\alpha, \alpha' \in \mathbb{N}^k}$  is defined as

$$q_{\alpha, \alpha'} := \max\{q_{l+1} : \text{for } l \text{ such that } \alpha(l) = \alpha'(l)\}.$$

The overlap matrix is clearly symmetric. Also, if  $\alpha(1) \neq \alpha'(1)$ , then  $q_{\alpha, \alpha'} = q_1$  and  $q_{\alpha, \alpha} = q_{k+1} = 1$ . In addition, the following inequality holds by definition

$$(3) \quad q_{\alpha, \alpha'} \geq \min\{q_{\alpha, \alpha''}; q_{\alpha', \alpha''}\}$$

for any triplet  $\alpha, \alpha', \alpha''$ . For  $d_{\alpha, \alpha'} := 1 - q_{\alpha, \alpha'}$ , the inequality becomes the *ultrametric inequality*:  $d_{\alpha, \alpha'} \leq \max\{d_{\alpha, \alpha''}, d_{\alpha', \alpha''}\}$ . It implies that at least two overlaps in the triplet must be the same, and the distinct one, if any, must be greater than the redundant overlap. As we will see later in the construction of the cavity field, the overlap matrix is also the covariance matrix of a gaussian field labeled by  $\alpha \in \mathbb{N}^k$ . In particular, it is positive definite.

**Definition (GREM).** Let  $x \in \mathcal{M}_a^{<1}$  with atoms at  $\{q_l\}_{1 \leq l \leq k}$  and  $q_{k+1} = 1$  with  $x(q_l) = x_l$  and  $x(q_{k+1}) = 1$ . A GREM( $x$ ) is the pair  $(\xi, Q)$  where  $\xi = \{\xi_\alpha\}_{\alpha \in \mathbb{N}^k}$  is the resulting point process constructed above with parameters  $x_l$ ,  $1 \leq l \leq k$ , and  $Q = \{q_{\alpha, \alpha'}\}$  is the corresponding symmetric, positive definite matrix with  $q_{\alpha, \alpha'} \in \{q_1, \dots, q_k, 1\}$ . Note that, by definition, a GREM( $x$ ) is a ROST.

We stress that the letter  $\xi$  will be used for the resulting process  $\xi = \{\xi_\alpha\}_{\alpha \in \mathbb{N}^k}$  and not for the whole cascade of processes. In particular,  $\xi$  does not contain information on the hierarchy. The information on the hierarchy is encoded in the labeling  $\alpha$  and expressed through the overlap matrix  $Q$ .

The point process  $\xi = \{\xi_\alpha\}_{\alpha \in \mathbb{N}^k}$  keeps some regularity features of the REM. Indeed,  $\sum_\alpha \xi_\alpha < \infty$  a.s. and  $\xi$  is a random Poisson process whose intensity measure, conditioned on

$\mathcal{F}_{k-1}^\xi$ , is  $\sum_{\alpha(k-1) \in \mathbb{N}^{k-1}} \xi_{\alpha(k-1)}^{x_k} x_k s^{x_k-1} ds$ . These facts are consequences of basic properties of the REM. Proofs can be found in [4], Lemma 2.1.

As for the REM, the summability allows the enumeration of the points of a realization in decreasing order, i.e.  $\{\xi_\alpha\}_{\alpha \in \mathbb{N}^k} = \{\xi_i\}_{i \in \mathbb{N}}$ ,  $\xi_1 > \xi_2 > \dots$ . This ordering induces a random bijection  $\phi : \mathbb{N} \rightarrow \mathbb{N}^k$  where  $\phi(i) = \alpha$  if  $\xi_\alpha = \xi_i$ . The matrix  $\phi^{-1} \circ Q \circ \phi = \{q_{\phi(i)\phi(j)}\}_{i,j \in \mathbb{N}}$  is now clearly random. We will sometimes abuse notation and write  $q_{ij}$  for  $q_{\phi(i)\phi(j)}$  and  $Q$  for  $\phi^{-1} \circ Q \circ \phi$ . The intended meaning will be clear from the notation and the context. The distribution of the matrix  $\phi^{-1} \circ Q \circ \phi$  looks intricate at first due to its dependence on the ordering of the process  $\xi$ . It turns out it has a simple form due to Bolthausen and Sznitman [4].

To illustrate this distribution, we must define the following random equivalence relations on  $\mathbb{N}$  for each level  $l$ ,  $0 \leq l \leq k$ :

$$(4) \quad i \sim_l j \text{ if and only if } [\phi(i)](l) = [\phi(j)](l).$$

By convention,  $i \sim_0 j$  for all  $i, j \in \mathbb{N}$ . We write  $\Gamma_{x_l}$  for the partition of  $\mathbb{N}$  induced by  $\sim_l$ , i.e.  $\Gamma_{x_l} := \mathbb{N} / \sim_l$ .  $\Gamma_{x_l}$  is obtained by lumping equivalence classes of  $\Gamma_{x_{l+1}}$ . This is because  $i \sim_l j$  if  $i \sim_{l+1} j$  by the definition (4). The distribution of the sequence of partitions  $\{\Gamma_{x_l}\}$  is surprisingly simple.

**Theorem 2** (Theorems 1.2, 2.2 and Proposition 1.4 in [4]). *Define  $\Gamma(t) := \Gamma_{e^{-t}}$ . The process  $\Gamma(t)$ ,  $t = 0, -\log x_k, \dots, -\log x_1 < \infty$ , is a discrete-time Markov process on the space of partitions of  $\mathbb{N}$  whose transition probabilities are defined as follows.*

*Consider  $\Gamma^{(n)}(s)$  and  $\Gamma^{(n)}(t)$ , sets of equivalence classes of  $\{1, \dots, n\}$ . Define  $k_s = |\Gamma^{(n)}(s)|$  and  $k_t = |\Gamma^{(n)}(t)|$ . Let  $\Gamma^{(n)}(t)$  be obtained from  $\Gamma^{(n)}(s)$  by respectively lumping  $m_1, \dots, m_{k_t}$  classes of  $\Gamma^{(n)}(s)$ . Then the transition probability from  $\Gamma^{(n)}(s)$  to  $\Gamma^{(n)}(t)$ ,  $s < t$ , is*

$$P(s \rightarrow t; \Gamma^{(n)}(s), \Gamma^{(n)}(t)) = \frac{(k_t - 1)!}{(k_s - 1)!} \left( \frac{e^{-t}}{e^{-s}} \right)^{(k_t-1)} \prod_{l=1}^{k_t} u(m_l, e^{-t}/e^{-s})$$

where  $u(1, x) = 1$  and  $u(m, x) = \frac{(m-1-x) \dots (1-x)}{(m-1)!}$ .

Moreover, the process  $\Gamma(t)$  is independent from the normalized point process  $\{\xi_i / \sum_i \xi_i\}_{i \in \mathbb{N}}$ .

Given a realization of the equivalence relations  $\sim_l$ , the overlap of the  $i$ -th and  $j$ -th points follows from equation (4):

$$q_{ij} := q_{\phi(i)\phi(j)} = q_{\max\{l+1: i \sim_l j\}}$$

Clearly, the distribution of the process  $Q = \{q_{ij}\}$  on symmetric, positive definite matrices depends uniquely on the distribution of the random equivalence classes given above. As an example of calculation of overlap probabilities using Theorem 2, we have  $\mathbb{P}_x(q_{12} = q_l) = (x_l - x_{l-1})$ . It is important to note that the last assertion of Theorem 2 implies that the process  $Q$  is independent from the normalized weights  $\{\xi_i / \sum_i \xi_i\}_{i \in \mathbb{N}}$ .

**2.4. The Cavity Field of the GREM.** In this section, we introduce the gaussian fields on the GREM that will appear in the definition of the functionals of interest on RPC's.

**Definition** (Gaussian Field on the GREM). *Let  $(\xi, Q)$  be a  $GREM(x)$  and  $g : [0, 1] \rightarrow \mathbb{R}^+$ , a strictly increasing function in  $C^1([0, 1])$  with  $g(0) = 0$ . A gaussian field with covariance function  $g$  on  $\xi$  is a centered gaussian process  $\kappa = (\kappa_\alpha(r), \alpha \in \mathbb{N}^k, 0 \leq r \leq 1)$  with*

covariance

$$\text{Cov}(\kappa_\alpha(q), \kappa_{\alpha'}(q')) = \int_0^{q \wedge q' \wedge q_{\alpha, \alpha'}} g'(r) dr = g(q \wedge q' \wedge q_{\alpha, \alpha'}).$$

In the case  $g(q) = q$ , the field is called the cavity field of  $\xi$  and will be denoted by  $\eta$ . We will write  $\kappa_i$  for  $\kappa_{\phi(i)}$ , the field of the  $i$ -th point of the configuration.

The cavity field  $\eta$  can be constructed explicitly. We consider, for each  $\alpha(l) \in \mathbb{N}^l$  and each  $0 \leq l \leq k$ , independent standard brownian motions  $B_{\alpha(l)}$  on  $[q_l, q_{l+1})$ . One could think of the brownian motion as attached to each subbranch  $\alpha(l)$  of the cascade. For each  $\alpha \in \mathbb{N}^k$ , we construct the process  $\eta_\alpha$  recursively. We set  $\eta_{\alpha(0)}(q) = B_{\alpha(0)}$  for  $0 \leq q \leq q_1$  and  $\eta_{\alpha(l)}(q) = \eta_{\alpha(l-1)}(q_l) + B_{\alpha(l)}(q)$  for  $q_l \leq q \leq q_{l+1}$ . Finally,  $\eta_\alpha(q) = \eta_{\alpha(l)}(q)$  where  $q_l \leq q \leq q_{l+1}$ . It is straightforward to check that  $\text{Cov}(\eta_\alpha(q), \eta_{\alpha'}(q')) = q \wedge q' \wedge q_{\alpha, \alpha'}$ . In particular,  $\text{Cov}(\eta_\alpha(1), \eta_{\alpha'}(1)) = q_{\alpha, \alpha'}$  which shows that the overlap matrix  $Q$  is positive definite as claimed before. The gaussian field  $\kappa$  with covariance function  $g$  can now be written as a stochastic integral on  $\eta$

$$\kappa_\alpha(q) = \int_0^q \sqrt{g'(r)} d\eta_\alpha(r).$$

The natural filtration of the cavity field on the tree is  $\mathcal{F}_q^\eta = \sigma(\eta_\alpha(r), \alpha \in \mathbb{N}^k, 0 \leq r \leq q)$ . It is useful to construct the filtration of a  $\text{GREM}(x)$  which keeps track of the information of the cascade of point processes as well as the cavity field on it. Let  $\mathcal{F} = (\mathcal{F}_q, q \in [0, 1])$  be the right-continuous filtration defined by  $\mathcal{F}_q = \mathcal{F}_{l-1}^\xi \otimes \mathcal{F}_q^\eta$  for  $q \in [q_l, q_{l+1})$  and for  $q = 1$ ,  $\mathcal{F}_1 = \mathcal{F}_k^\xi \otimes \mathcal{F}_1^\eta$ . Note that the filtration is actually continuous at every point except at the points  $q_l$  where branchings occur. Also, the  $\sigma$ -algebra  $\mathcal{F}_q$  contains the information of the branchings strictly above  $[q_l, q_{l+1})$  for  $q$  in this interval as symbolized by  $\mathcal{F}_{l-1}^\xi$ .

**Definition.** Throughout the rest of this paper, we will write  $\mathbb{P}_x$  for the probability measure on  $\mathcal{F}_1$  of a  $\text{GREM}(x)$  and its cavity field.  $\mathbb{E}_x$  will denote the expectation.

### 3. THE QUASI-STATIONARITY PROPERTY OF THE GREM

The quasi-stationarity property of the REM stated in Proposition 1 induces a similar stability property on a cascade of REM. We now study this important feature of the GREM process. The stochastic shift at each point will be written as a function of the gaussian field on the RPC presented in the last section. We also look at the distribution of the field after a shift. It is modified by the reordering as it was in the REM case.

We first need to introduce a class of function for which the stochastic shift is well-defined.

**Definition.** Let  $\mathcal{C}$  be the class of functions  $\psi$  in  $C^2(\mathbb{R})$  satisfying

- $\psi'$  and  $\psi''$  are bounded on  $\mathbb{R}$ ;
- $E_Y [e^{\psi(Y)}] < \infty$  for any gaussian variable  $Y$ .

We remark that  $E_Y [e^{x\psi(Y)}] < \infty$  for any  $0 < x < 1$  (using Jensen's inequality applied with the convex function  $f(y) = y^{1/x}$ ). Also,  $\mathcal{C}$  includes the functions  $\psi(\kappa) = \log \cosh(\beta\kappa + h)$  and  $\psi(\kappa) = \beta\kappa$ .

In this section, we are interested in the stability properties of the GREM under the stochastic shift

$$\xi_\alpha \mapsto \xi_\alpha e^{f_q(\kappa_\alpha(q))}$$

where  $\kappa$  is a gaussian field on  $\xi$  with covariance function  $g$ . The family of functions  $(f_q, q \in [0, 1])$  is assumed to be contained in the class  $\mathcal{C}$ . We also assume that the family is differentiable, i.e. for  $y$  fixed,  $f_q(y)$  is a differentiable function of  $q$ .

An example of such a family is the following. Fix  $x \in \mathcal{M}_a^{<1}$ . Let  $\psi \in \mathcal{C}$  and  $\psi_1 := \psi$ . For  $q \in [q_l, q_{l+1})$ , we define recursively  $\psi_q$  as

$$(5) \quad \begin{aligned} \psi_q(y) &= \frac{1}{x_l} \log E_z \left[ e^{x_l \psi_{q_{l+1}}(y+z\sqrt{g(q_{l+1})-g(q)})} \right] \text{ for } 1 \leq l \leq k; \\ \psi_q(y) &= \log E_z \left[ e^{\psi_{q_{l+1}}(y+z\sqrt{g(q_1)-g(q)})} \right] \text{ for } l = 0 \end{aligned}$$

where  $E_z$  denotes the expectation over  $z$ , a standard gaussian. We dropped the dependence of  $\psi_q$  on  $x$  and  $g$  but the reader must keep in mind this dependence. It is easy to verify that  $\psi_q$  is in  $\mathcal{C}$ . Moreover, by direct derivation,  $\psi_q$  is seen to satisfy the differential equation (see e.g. [13])

$$\partial_q \psi_q(y) + \frac{g'(q)}{2} (\psi_q''(y) + x_l (\psi_q'(y))^2) = 0$$

with the continuity condition  $\lim_{q \rightarrow q_{l+1}^-} \psi_q(y) = \psi_{q_{l+1}}(y)$ . More generally, by using the condition of continuity and the equation for each interval, the function  $\psi_q(y)$  actually satisfies

$$\partial_q \psi_q(y) + \frac{g'(q)}{2} (\psi_q''(y) + x(q) (\psi_q'(y))^2) = 0$$

with boundary condition  $\psi_1(y) = \psi(y)$ .

As a matter of fact, the function  $e^{\psi_q(y)}$  is the factor  $E[W^{x_l}]^{1/x_l}$  coming from the application of the quasi-stationarity property of Proposition 1 to a REM( $x_l$ ) with a shift  $W_\alpha = e^{\psi_r(y+Y_\alpha)}$ ,  $q < r \leq q_{l+1}$ , where  $Y_\alpha$  are independent gaussians  $\mathcal{N}(0, g(r) - g(q))$ .

The family of functions  $\psi_q$  has the property that the GREM distribution under their associated stochastic shift is invariant up to a random common factor (labeled by  $\alpha(0)$ ):

$$(6) \quad (\xi_\alpha e^{\psi_q(\kappa_\alpha(q))}, Q) \stackrel{\mathcal{D}}{=} (\xi_\alpha e^{\psi_{q_1}(\kappa_{\alpha(0)}(q_1))}, Q)$$

for  $q \in [q_1, 1]$ . Indeed, let  $q \in [q_l, q_{l+1}]$ . We write  $\delta_{\alpha(l)}(q) := \kappa_{\alpha(l)}(q) - \kappa_{\alpha(l)}(q_l)$ . Note that the  $\delta_{\alpha(l)}$ 's are independent gaussian  $\mathcal{N}(0, g(q) - g(q_l))$  for each  $\alpha(l)$ . Therefore, we can apply the quasi-stationarity to each REM  $\zeta^{\alpha(l-1)}$ :

$$\begin{aligned} \xi_\alpha e^{\psi_q(\kappa_\alpha(q))} &= \xi_{\alpha(l-1)} \left( \zeta^{\alpha(l-1)} \right)_{\alpha_l} e^{\psi_q(\kappa_\alpha(q_l) + \delta_{\alpha(l)}(q))} \prod_{l'=l+1}^k \left( \zeta^{\alpha(l'-1)} \right)_{\alpha'_l} \\ &\stackrel{\mathcal{D}}{=} \xi_{\alpha(l-1)} \left( \zeta^{\alpha(l-1)} \right)_{\alpha_l} E_{\delta_{\alpha(l)}} \left[ e^{x_l \psi_q(\kappa_\alpha(q_l) + \delta_{\alpha(l)}(q))} \right]^{1/x_l} \prod_{l'=l+1}^k \left( \zeta^{\alpha(l'-1)} \right)_{\alpha'_l} \\ &= \xi_\alpha e^{\psi_{q_l}(\kappa_\alpha(q_l))} \end{aligned}$$

where we used the definition of  $\psi_{q_l}$  in the last equality. This procedure is applied successively up to  $q_1$  to prove the claim.

Surprisingly, it turns out that the property (6) characterizes the family  $(\psi_q, q \in [0, 1])$ .

**Theorem 3** (Quasi-Stationarity of the GREM). *Let  $(f_q, q \in [0, 1])$  be a differentiable family of functions in  $\mathcal{C}$ . Let  $(\xi, Q)$  be a GREM( $x$ ) and  $\kappa$ , its gaussian field with covariance*

function  $g$ . Then the function  $f_q(y)$  satisfies the differential equation (3) for  $q \in [q_1, 1]$  if and only if

$$(\xi_\alpha e^{f_q(\kappa_\alpha(q))}, Q) \stackrel{\mathcal{D}}{=} (\xi_\alpha e^{f_r(\kappa_\alpha(r))}, Q)$$

for all  $q, r \in [q_1, 1]$ . In particular,

$$(7) \quad (\xi_\alpha e^{f_1(\kappa_\alpha(1))}, Q) \stackrel{\mathcal{D}}{=} (\xi_\alpha e^{f_{q_1}(\kappa_{\alpha(0)}(q_1))}, Q).$$

*Proof.* The sufficiency of the differential equation for stability was proven above. We prove the necessity. Pick  $q \in [q_l, q_{l+1})$ ,  $l \geq 1$ . Choose  $\Delta q$  small enough so that  $q + \Delta q$  belongs also to  $[q_l, q_{l+1})$ . Consider  $\delta_{\alpha(l)}(\Delta q) := \kappa_{\alpha(l)}(q + \Delta q) - \kappa_{\alpha(l)}(q)$  that are independent gaussian  $\mathcal{N}(0, g(q + \Delta q) - g(q))$  for each  $\alpha(l)$ . Applying quasi-stationarity, we obtain

$$\xi_\alpha e^{f_{q+\Delta q}(\kappa_\alpha(q) + \delta_{\alpha(l)}(\Delta q))} \stackrel{\mathcal{D}}{=} \xi_\alpha E_{\delta_{\alpha(l)}} \left[ e^{x_l f_{q+\Delta q}(\kappa_\alpha(q) + \delta_{\alpha(l)}(\Delta q))} \right]^{1/x_l}.$$

But, by equation (3), we also have

$$\xi_\alpha e^{f_{q+\Delta q}(\kappa_\alpha(q) + \delta_{\alpha(l)}(\Delta q))} \stackrel{\mathcal{D}}{=} \xi_\alpha e^{f_q(\kappa_\alpha(q))}.$$

As this must hold for all realization of  $\kappa_\alpha(q)$ , we conclude that  $E_{\delta_{\alpha(l)}} \left[ e^{x_l f_{q+\Delta q}(y + \delta_{\alpha(l)}(\Delta q))} \right] = e^{x_l f_q(y)}$  for all  $\Delta q$ . In particular,

$$\lim_{\Delta q \rightarrow 0} \frac{1}{\Delta q} \left( E_{\delta_{\alpha(l)}} \left[ e^{x_l f_{q+\Delta q}(y + \delta_{\alpha(l)}(\Delta q))} \right] - e^{x_l f_q(y)} \right) = 0$$

if the limit exists. The limit does exist and is easily computed by Itô's formula

$$\left. \frac{d}{dr} E_{\delta_{\alpha(l)}} \left[ e^{x_l f_r(y + \delta_{\alpha(l)}(r))} \right] \right|_{r=0} = x_l e^{x_l f_q(y)} \left( \partial_q f_q(y) + \frac{g'(q)}{2} (f_q''(y) + x_l (f_q'(y))^2) \right) = 0$$

which yields the desired differential equation.  $\square$

From now on, we will use the notation  $\tilde{\xi}_\alpha$  for the shifted process  $\xi_\alpha e^{\psi(\kappa_\alpha(1))}$ . Consider the random permutation  $\pi$  of  $\mathbb{N}$  induced by this random shift, i.e.  $\pi(i) = i'$  if  $\tilde{\xi}_{i'} = \xi_i e^{\psi(\kappa_i(1))}$ . As in Proposition 1, we are interested in the backward distribution of the random shift. More precisely, we study the distribution of the field  $\tilde{\kappa}$ :

$$\tilde{\kappa} := (\tilde{\kappa}_i, i \in \mathbb{N}) = (\kappa_{\pi^{-1}(i)}, i \in \mathbb{N}).$$

**Proposition 4** (Backward Distribution of the field). *Let  $\mu_Q^{(n)}$  be the joint distribution of the gaussian fields  $\kappa_1, \dots, \kappa_n$  under  $\mathbb{P}_x$  given the overlap  $Q$  and  $\tilde{\mu}_Q^{(n)}$ , the distribution of  $\tilde{\kappa}_1, \dots, \tilde{\kappa}_n$  under  $\mathbb{P}_x$  given  $Q$ . Then  $\tilde{\mu}_Q^{(n)}$  is absolutely continuous with respect to  $\mu_Q^{(n)}$  and*

$$\frac{d\tilde{\mu}_Q^{(n)}}{d\mu_Q^{(n)}} = \prod_{l=0}^k \prod_{i \in \{1, \dots, n\} / \sim_l} V_i$$

where the product is over the equivalence classes of  $\{1, \dots, n\} / \sim_l$  and

$$V_i = \frac{e^{x_l \psi_{q_{l+1}}(\tilde{\kappa}_i(q_{l+1}))}}{e^{x_l \psi_{q_l}(\tilde{\kappa}_i(q_l))}}.$$



From the form of the distribution, we conclude that  $\tilde{\kappa}$  is independent from  $\tilde{\xi}$ . The last proposition seems abstract at first but will prove extremely useful as it will allow us to express derivatives of the Parisi functional (and more generally, of functionals on RPC's) in a compact way as expectations over the backward gaussian field.

*Proof.* We prove the case  $n = 1$ . Choose a finite set of  $t_j$ 's in  $[0, 1]$ ,  $j = 1, \dots, J$ . We consider the union of  $\{t_j\}_{j \in J}$  and  $\{q_l\}_{1 \leq l \leq k+1}$ :  $\{r_i\}_{i=1, \dots, N} = \{t_j\} \cup \{q_l\}$  so  $r_1 < r_2 < \dots < r_N$ . The claim will be proven for  $n = 1$  if for any choice of  $t_j$ , the finite-dimensional distributions satisfy

$$\tilde{\mu}^{(1)}(\tilde{\kappa}_1(t_j) \in A_j, 1 \leq j \leq J) = \int_{B_1} \dots \int_{B_N} \prod_{l=0}^k \frac{e^{x_l \psi_{q_{l+1}}(\tilde{\kappa}_1(q_{l+1}))}}{e^{x_l \psi_{q_l}(\tilde{\kappa}_1(q_l))}} \mu^{(1)}(\tilde{\kappa}_1(r_i) \in ds_i, 1 \leq i \leq N).$$

where  $B_i = A_j$  if  $r_i = t_j$  and  $B_i = \mathbb{R}$  otherwise.

Define  $l_i$  as the level  $l$  where  $r_i \in (q_l, q_{l+1}]$ . By definition, if  $r_i \neq q_l$  for any  $l$  then  $l_{i+1} = l_i$ . The key point in the definition of the  $r_i$  is that the  $\kappa_{\alpha(l_i)}$ 's, given  $\mathcal{F}_{r_{i-1}}$ , are independent. (This would not be true if one does not consider all  $q_l$ 's in the definition of  $r_i$ 's.)

By Theorem 3, we have that,

$$\tilde{\xi}_{\alpha(l_i)} = \xi_{\alpha(l_i)} e^{\psi_{r_i}(\kappa_{\alpha(l_i)}(r_i))} \stackrel{\mathcal{D}}{=} \xi_{\alpha(l_i)} e^{\psi_{r_{i-1}}(\kappa_{\alpha(l_i)}(r_{i-1}))}.$$

By the quasi-stationarity property of the REM (Proposition 1), the backward distribution of  $\kappa_1(r_i)$  given  $\mathcal{F}_{r_{i-1}}$  is

$$\frac{e^{x_{l_i} \psi_{r_i}(\tilde{\kappa}_1(r_i))}}{e^{x_{l_i} \psi_{r_{i-1}}(\tilde{\kappa}_1(r_{i-1}))}} \mu^{(1)}(\tilde{\kappa}_1(r_i) \in ds | \mathcal{F}_{r_{i-1}}).$$

The joint distribution of the  $\tilde{\kappa}_1(r_i)$  is therefore

$$\tilde{\mu}^{(1)}(\tilde{\kappa}_1(r_i) \in B_i, 1 \leq i \leq N) = \int_{B_1} \dots \int_{B_N} \prod_{i=1}^N \frac{e^{x_{l_i} \psi_{r_i}(\tilde{\kappa}_1(r_i))}}{e^{x_{l_i} \psi_{r_{i-1}}(\tilde{\kappa}_1(r_{i-1}))}} \mu^{(1)}(\tilde{\kappa}_1(r_i) \in ds_i, 1 \leq i \leq N).$$

If  $r_i \neq q_l$  for any  $l$ , we noted that  $l_{i+1} = l_i$ . Thus, the factors in the product coming from these terms cancel out and only the terms in  $q_l$  remain. The claim is proven for  $n = 1$ .

If  $n > 1$ , the same procedure applies. For example, one could pick  $\{t_j^{(m)}\}_j$  where  $m$  indexes the points from 1 to  $n$ . Then, one considers  $\{r_i\} = \{q_l\} \cup_m \{t_j^{(m)}\}$  as before. In this case, at each level  $l$ , every equivalence class  $\bar{i}$  of  $\{1, \dots, n\} / \sim_l$  picks up a Radon-Nikodym derivative  $V_{\bar{i}}$ .  $\square$

We now state a result on the regularity of the expectations of the field  $\tilde{\kappa}$  that we will need when studying functionals of the GREM's. We omit the proof as it is a direct consequence of the fact that  $\tilde{\mu}_Q^{(n)}$  is smooth in the parameters  $x_l$ 's and  $q_l$ 's.

**Corollary 5** (Expectations of the Backward Field). *Let  $x \in \mathcal{M}_a^{<1}$ . Consider  $(\xi, Q)$ , a GREM( $x$ ), with a gaussian field  $\kappa$ . Let  $Q_n$  be an  $n \times n$  matrix such that  $\mathbb{P}_x(Q_n) \neq 0$  and define*

$$F_{Q_n}(\mathbf{x}, \mathbf{q}) = \mathbb{E}_x \left[ \prod_{i=1}^n \phi(\tilde{\kappa}_i(1)) \middle| Q_n \right]$$

where  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  is a bounded continuous function and  $\mathbf{q} = (q_0, \dots, q_k)$  and  $\mathbf{x} = (x_0, \dots, x_k)$  as before. The following statements hold :

- i*  $F_{Q_n}(\mathbf{x}, \mathbf{q})$  is a continuous function of  $\mathbf{q}$ ;
- ii*  $F_{Q_n}(\mathbf{x}, \mathbf{q})$  is continuous in  $x_j$  on  $(x_{j-1}, x_{j+1})$ , for all  $1 \leq j \leq k$ . The limits  $\lim_{x_j \rightarrow x_{j-1}} F_{Q_n}(\mathbf{x}, \mathbf{q})$  and  $\lim_{x_j \rightarrow x_{j+1}} F_{Q_n}(\mathbf{x}, \mathbf{q})$  exists and are continuous functions of  $\mathbf{q}$ .

In the last claim, we stress out that the limit of  $F_{Q_n}$  as  $x_j \rightarrow x_{j+1}$  is not equal to  $F_{Q_n}$  evaluated at  $(x_0, \dots, x_{j-1}, x_{j+1}, x_{j+1}, \dots, x_k)$  and similarly for the limit  $x_j \rightarrow x_{j-1}$ . This is basically because the distribution functions of the approximating sequence clearly possess one atom more than the limiting distribution function with  $x_j = x_{j+1}$ . Therefore,  $\lim_{x_j \rightarrow x_{j+1}} F_{Q_n}(\mathbf{x}, \mathbf{q})$  differs from the function  $F_{Q_n}$  evaluated at the limiting distribution function as the product over the number of atoms appearing in the density of  $\tilde{\kappa}$  in Proposition 4 has an extra factor in the first case.

#### 4. THE PARISI FUNCTIONAL AS A FUNCTIONAL ON RPC'S

In this section, we establish that functionals on RPC's of the form (2) coincide with Parisi-like functionals on the space  $\mathcal{M}_a^{<1}$ . The RPC formulation is a natural framework for deriving properties of such functionals as one can take advantage of the rich structure of the GREM and of the quasi-stationarity property. This is useful as it can be sometimes tedious to derive properties from the solution to the differential equation (1). As an example of techniques in the RPC formulation, we obtain differentiation formulas which lead to the continuity and the monotonicity of the functional. This result slightly generalizes a theorem of Guerra [6]. Using continuity, we also exhibit a limit form of the RPC functional in the singular case  $x \in \mathcal{M}_a$  but  $x \notin \mathcal{M}_a^{<1}$ .

**4.1. The Parisi Functional.** We first define the Parisi functional in a general way. The reader can consult [6, 10, 14] for particular choices of settings.

**Definition.** Let  $\psi \in \mathcal{C}$  and  $g \in C^1([0, 1])$ , a strictly increasing function such that  $g(0) = 0$ . The Parisi functional  $\mathcal{P}_{\psi, g}^{par} : \mathcal{M}_a \rightarrow \mathbb{R}$  is defined as  $\mathcal{P}_{\psi, g}^{par}(x) := f_x(0, 0)$  where  $f_x(q, y)$  is the solution to the differential equation

$$\partial_q f(q, y) + \frac{g'(q)}{2} \left[ \partial_y^2 f(q, y) + x(q) (\partial_y f(q, y))^2 \right] = 0$$

with the boundary condition  $f(1, y) = \psi(y)$ . The case  $\psi(\eta) = \log \cosh(\beta\eta + h)$  and  $g(q) = q$  reduces to equation (1).

From another perspective, the  $AS^2$  variational principle formulation is based on functionals of the following form [2, 3].

**Definition.** Let  $\psi$  and  $g$  be as above. The functional  $\mathcal{P}_{\psi, g} : \mathcal{M}_a^{<1} \rightarrow \mathbb{R}$  is defined by

$$(8) \quad \mathcal{P}_{\psi, g}(x) := \mathbb{E}_x \left[ \log \frac{\sum_{\alpha} \xi_{\alpha} e^{\psi(\kappa_{\alpha}(1))}}{\sum_{\alpha} \xi_{\alpha}} \right]$$

where  $(\xi, Q)$  is a GREM( $x$ ) and  $\kappa$  is a gaussian field on  $\xi$  with covariance  $\text{Cov}(\kappa_{\alpha}(q), \kappa_{\alpha'}(q')) = g(q \wedge q' \wedge q_{\alpha, \alpha'})$ . We will write  $\mathcal{P}_{\beta, h}$  in the special case  $\psi(\eta) = \log \cosh(\beta\eta + h)$  and  $g(q) = q$ .

The main result of the section is the correspondence of these functionals and the Parisi functionals.

**Theorem 6** (The Parisi functional on RPC's). *For any  $x \in \mathcal{M}_a^{<1}$ ,  $\mathcal{P}_{\psi, g}^{par}(x) = \mathcal{P}_{\psi, g}(x)$ .*

*Proof.* The quasi-stationarity of the GREM (Theorem 3) implied equation (7):

$$\xi_\alpha e^{\psi(\kappa_\alpha(1))} \stackrel{\mathcal{D}}{=} \xi_\alpha e^{\psi_{q_1}(\kappa_{\alpha(0)}(q_1))}.$$

By inserting this into equation (8), one gets

$$\mathcal{P}_{\psi,g}(x) = \mathbb{E}_x [\psi_{q_1}(\kappa_{\alpha(0)}(q_1))] = \psi_0(0)$$

by definition of  $\psi_0$  (equation (5)). But  $\psi_q(y)$  is the solution to the differential equation of the Parisi functional by Theorem 3 again. We conclude that  $\mathcal{P}_{\psi,g}(x) = \mathcal{P}_{\psi,g}^{par}(x)$ .  $\square$

It is worth pointing out that  $\mathcal{P}_{\psi,g}$  is linear when  $\psi$  is.

**Proposition 7** (Linearity of  $\mathcal{P}_{\psi,g}$ ). *If  $\psi(\eta) = \beta\eta$ , then  $\mathcal{P}_{\psi,g}(x)$  is a bounded linear functional. Precisely,  $\mathcal{P}_{\psi,g}(x) = \frac{\beta^2}{2} \int_0^1 x(q)dg(q)$ .*

*Proof.* We can write  $\kappa_\alpha(1)$  as a sum of independent increments:  $\kappa_\alpha(1) = \sum_{l=0}^k \delta_{\alpha(l)}$  where  $\delta_{\alpha(l)} = \kappa_\alpha(q_{l+1}) - \kappa_\alpha(q_l)$ . We can then apply the quasi-stationarity to each REM  $\zeta^{\alpha(l-1)}$  of the cascade:

$$\xi_\alpha e^{\beta \sum_{l=0}^k \delta_{\alpha(l)}} \stackrel{\mathcal{D}}{=} e^{\beta \kappa_{\alpha(0)}(q_1)} \xi_\alpha \prod_{l=1}^k e^{\frac{\beta^2}{2} x_l(g(q_{l+1}) - g(q_l))}$$

where we have used the fact that the Laplace transform of a standard gaussian is  $e^{\frac{\lambda^2}{2}}$ . Note that the expectation of  $\kappa_{\alpha(0)}(q_1)$  is 0 by definition. The claim is proven by inserting the last equation into equation (8) for  $\mathcal{P}_{\psi,g}$ .  $\square$

**4.2. Differentiation.** How does  $\mathcal{P}_{\psi,g}$  vary as the position  $q_l$  of an atom or the weight  $x(q_l)$  is changed? To address this question, we look at the derivatives of  $\mathcal{P}_{\psi,g}(x)$  with respect to the parameters  $x_l$  and  $q_l$ . It turns out that the  $x$ -derivatives are simply related to the  $q$ -derivatives due to the structure of the space  $\mathcal{M}_a$ . Both are expressed as expectations on the backward field  $\tilde{\kappa}$ . For conciseness, we will use the notation  $\kappa_i \equiv \kappa_i(1)$  for a gaussian field  $\kappa_i$  at  $q = 1$  throughout the section.

**Proposition 8** ( $q$ -Derivatives). *Let  $x \in \mathcal{M}_a^{<1}$  with atoms  $\{q_l\}$ . The following differentiation formulas hold:*

$$(9) \quad \partial_{q_l} \mathcal{P}_{\psi,g}(x) = \frac{-g'(q_l)}{2} \mathbb{E}_x [\psi'(\tilde{\kappa}_1) \psi'(\tilde{\kappa}_2) \chi_{q_{12}=q_l}]$$

where  $\kappa$  is a gaussian field on a GREM( $x$ ) and  $\chi_A$  is the identity function of the event  $A$ . If  $g(q) = q$ ,

$$(10) \quad \begin{aligned} \partial_{q_i} \partial_{q_j} \mathcal{P}_{\psi,g}(x) = & -\frac{3}{2} \mathbb{E}_x \left[ \prod_{m=1}^4 \psi'(\tilde{\eta}_m) \chi_{q_{12}=q_i, q_{34}=q_j} \right] \\ & + 2 \mathbb{E}_x [\psi'(\tilde{\eta}_1) (\psi''(\tilde{\eta}_2) + \psi'^2(\tilde{\eta}_2)) \psi'(\tilde{\eta}_3) \chi_{q_{12}=q_i, q_{23}=q_j}] \\ & - \frac{1}{2} \delta_{ij} \mathbb{E}_x [(\psi''(\tilde{\eta}_1) + \psi'^2(\tilde{\eta}_1)) (\psi''(\tilde{\eta}_2) + \psi'^2(\tilde{\eta}_2)) \chi_{q_{12}=q_i}] \end{aligned}$$

where  $\eta$  is the cavity field of the GREM( $x$ ). In particular, both derivatives are continuous functions of the  $q_l$ 's.

*Proof.* The continuity of the derivatives follows from the form of the expression and Corollary 5. For the first derivative, we use the gaussian differentiation formula (25) in Appendix B. The parameter is  $q_l$  and the covariance is  $\mathbb{E}_x[\kappa_\alpha(1)\kappa_{\alpha'}(1)] = \sum_{l=1}^{k+1} g(q_l)\delta_{q_{\alpha,\alpha'},q_l}$  where  $\delta_{q,q'} = 1$  if  $q = q'$  and 0 otherwise. Therefore  $\partial_{q_l}\mathbb{E}_x[\kappa_\alpha(1)\kappa_{\alpha'}(1)] = g'(q_l)\delta_{q_{\alpha,\alpha'},q_l}$  for  $l = 1, \dots, k$ . Note that the first term of formula (25) vanishes as  $q_{\alpha,\alpha} = 1$  for all  $\alpha$ . Thus, the first derivative becomes

$$\partial_{q_l}\mathcal{P}_{\psi,g}(x) = \frac{-g'(q_l)}{2}\mathbb{E}_x\left[\frac{\sum_{\alpha,\alpha'}\tilde{\xi}_\alpha\tilde{\xi}_{\alpha'}\psi'(\kappa_\alpha)\psi'(\kappa_{\alpha'})\delta_{q_{\alpha,\alpha'},q_l}}{\sum_{\alpha,\alpha'}\tilde{\xi}_\alpha\tilde{\xi}_{\alpha'}}\right]$$

where we have used the notation  $\tilde{\xi}_\alpha = \xi_\alpha e^{\psi(\kappa_\alpha(1))}$ . After ordering the  $\tilde{\xi}_\alpha$ , we can write by the definition of  $\tilde{\kappa}$

$$\partial_{q_l}\mathcal{P}_{\psi,g}(x) = \frac{-g'(q_l)}{2}\mathbb{E}_x\left[\frac{\sum_{i,j}\tilde{\xi}_i\tilde{\xi}_j\psi'(\tilde{\kappa}_i)\psi'(\tilde{\kappa}_j)\delta_{q_{ij},q_l}}{\sum_{i,j}\tilde{\xi}_i\tilde{\xi}_j}\right].$$

Conditioning on  $\tilde{\xi}$  and  $q_{ij}$ , bearing in mind the independence of  $\tilde{\kappa}$  and  $\tilde{\xi}$  from Proposition 4, one can write the expectation in the r.h.s. as

$$\mathbb{E}_x\left[\psi'(\tilde{\kappa}_1)\psi'(\tilde{\kappa}_2)|q_{12} = q_l\right]\mathbb{E}_x\left[\frac{\sum_{i,j}\tilde{\xi}_i\tilde{\xi}_j\delta_{q_{ij},q_l}}{\sum_{i,j}\tilde{\xi}_i\tilde{\xi}_j}\right].$$

where we have used the fact that the joint distribution of  $\tilde{\kappa}_i$  and  $\tilde{\kappa}_j$  is the same as the distribution of  $\tilde{\kappa}_1$  and  $\tilde{\kappa}_2$  for all  $i, j$ . The proof is completed by first recalling that the shifted normalized process  $\tilde{\xi}_i/\sum_i\tilde{\xi}_i$  is distributed as the original normalized process by the quasi-stationarity property. Moreover, the normalized process is independent of the overlap matrix  $\{q_{ij}\}$  as it was remarked in Theorem 2. Therefore, the second expectation above simply becomes  $\mathbb{P}_x(q_{12} = q_l)$  by conditioning on the normalized weights and by noticing that  $\mathbb{P}_x(q_{ij} = q_l)$  does not depend on  $i, j$  for  $i \neq j$ . This yields equation (9).

The second expression is obtained from formula (26) in Appendix B with  $\phi(\kappa) = \log \sum_\alpha \xi_\alpha e^{\psi(\kappa_\alpha)}$  and straightforward derivation.  $\square$

Obviously, a formula holds also for a more general  $g$  in the case of the second derivatives. We omit it for the sake of conciseness.

The nice feature of the above derivative formulas is that they explicitly express the derivatives of the Parisi functional in terms of expectations over the field  $\tilde{\kappa}$  and probabilities on the GREM cascade. Using conditional expectations, we may rewrite the second derivative formula, e.g. in the case  $\psi(z) = \log \cosh(\beta z + h)$  where the identity  $\psi'' + \psi'^2 = 1$  holds, as

(11)

$$\begin{aligned}\partial_{q_i}\partial_{q_j}\mathcal{P}_{\beta,h}(x) &= -\frac{3}{2}\beta^4 \sum_{Q_4(q_i,q_j)} \mathbb{P}_x(Q_4(q_i,q_j))\mathbb{E}_x\left[\prod_{m=1}^4 \tanh(\beta\tilde{\eta}_m + h)|Q_4(q_i,q_j)\right] \\ &\quad + 2\beta^4 \sum_{Q_3(q_i,q_j)} \mathbb{P}_x(Q_3(q_i,q_j))\mathbb{E}_x[\tanh(\beta\tilde{\eta}_1 + h)\tanh(\beta\tilde{\eta}_2 + h)|Q_3(q_i,q_j)] \\ &\quad - \frac{1}{2}\beta^4\delta_{ij}\mathbb{P}_x(q_{12} = q_i).\end{aligned}$$

where the sums are over all  $3 \times 3$  and  $4 \times 4$  matrices  $Q_3$  and  $Q_4$  such that  $q_{12} = q_i, q_{23} = q_j$  and  $q_{12} = q_i, q_{34} = q_j$  with  $\mathbb{P}_x(Q_3(q_i, q_j)) \neq 0$  and  $\mathbb{P}_x(Q_4(q_i, q_j)) \neq 0$ . Note in particular that these matrices must satisfy the inequality (3).

**Corollary 9** (Useful estimate on the  $q$ -derivatives). *Let  $x \in \mathcal{M}_a^{<1}$  with  $i$ -th atom at  $q_i$ . Let  $C > 0$  be such that  $|\psi'| \leq C$ . Then*

$$(12) \quad 0 \leq -\partial_{q_i} \mathcal{P}_{\psi, g}(x) \leq \frac{C^2}{2} (x_i - x_{i-1}) g'(q_i).$$

*Proof.* The upper bound is clear from equation (9), the assumption on  $\psi$  and the fact that  $\mathbb{P}_x(q_{12} = q_i) = x_i - x_{i-1}$ . The lower bound is a consequence of

$$\mathbb{E}_x \left[ \psi'(\tilde{\eta}_1) \psi'(\tilde{\eta}_2) \middle| q_{12} = q_i \right] = \mathbb{E}_x \left[ \mathbb{E}_x[\psi'(\tilde{\eta}_1) | \mathcal{F}_{q_i}]^2 \right]$$

□

The structure of the space  $\mathcal{M}_a$  enables us to relate the  $x$ -derivatives and the  $q$ -derivatives through a simple differentiation scheme. To do so, we need to define the measure obtained from  $x \in \mathcal{M}_a$  by transporting a mass  $\delta$  from an atom at  $q_i$  to  $s \in [0, 1]$ .

**Definition.** *Let  $x \in \mathcal{M}_a$ . Let  $q_i$  be the position of the  $i$ -th atom of  $x$ ,  $i \leq k$ . Consider  $\delta > 0$  and  $s \in [0, 1]$ . We define  $x_{\delta, q_i \rightarrow s} \in \mathcal{M}_a$  as the atomic measure obtained from  $x$  by transporting a mass  $\delta$  from  $q_i$  to  $s$ . Note that if  $s = q_j$  for some  $j$ , then  $x_{\delta, q_i \rightarrow s}$  has  $k$  atoms; otherwise,  $x_{\delta, q_i \rightarrow s}$  has  $k + 1$  atoms.*

**Proposition 10** ( $x$ -Derivatives). *Consider  $x \in \mathcal{M}_a^{<1}$ . Then*

$$(13) \quad \partial_{x_i} \mathcal{P}_{\psi, g}(x) = \frac{1}{2} \int_{q_i}^{q_{i+1}} \lim_{\delta \rightarrow 0} \mathbb{E}_{x_{\delta, q_{i+1} \rightarrow r}} \left[ \psi'(\tilde{\kappa}_1) \psi'(\tilde{\kappa}_2) \middle| q_{12} = r \right] dg(r).$$

*In particular, if  $|\psi'| \leq C$  for some  $C > 0$*

$$0 \leq \partial_{x_i} \mathcal{P}_{\psi, g}(x) \leq \frac{C^2}{2} (g(q_{i+1}) - g(q_i)).$$

*Proof.* The upper and lower bounds are obtained as in Corollary 9 followed by integration.

For the sake of clarity, we set  $x_\delta = x_{\delta, q_{i+1} \rightarrow r}$  throughout the proof. By Proposition 8,

$$-\partial_r \mathcal{P}_{\psi, g}(x_\delta) = \delta \frac{1}{2} g'(r) \mathbb{E}_{x_\delta} \left[ \prod_{m=1}^2 \psi'(\tilde{\kappa}_m) \middle| q_{12} = r \right].$$

Here we have used  $\mathbb{P}_{x_\delta}(q_{12} = r) = x_\delta(r) - x_\delta(q_i) = \delta$ . The r.h.s. of the above equation satisfies the fundamental theorem of calculus as a function of  $r$  as  $g'(r)$  and  $\mathbb{E}_{x_\delta} \left[ \prod_{i=1}^2 \psi'(\tilde{\kappa}_i) \middle| q_{12} = r \right]$  are continuous functions of  $r$  by definition and Corollary 5 respectively. Moreover, the limit  $\lim_{\delta \rightarrow 0} \frac{1}{\delta} \partial_r \mathcal{P}_{\psi, g}(x_\delta)$  exists and is bounded thanks to Lemma 5:

$$\lim_{\delta \rightarrow 0} \frac{1}{\delta} \partial_r \mathcal{P}_{\psi, g}(x_\delta) = \frac{1}{2} g'(r) \lim_{\delta \rightarrow 0} \mathbb{E}_{x_\delta} \left[ \prod_{m=1}^2 \psi'(\tilde{\kappa}_m) \middle| q_{12} = r \right].$$

The fundamental theorem of calculus and the above limit yield the desired expression for the right-derivative  $\partial_{x_i}^+ \mathcal{P}_{\psi,g}(x)$ :

$$\begin{aligned} \partial_{x_i}^+ \mathcal{P}_{\psi,g}(x) &= \lim_{\delta \rightarrow 0} \frac{\mathcal{P}_{\psi,g}(x_{\delta,q_{i+1} \rightarrow q_i}) - \mathcal{P}_{\psi,g}(x)}{\delta} \\ &= \lim_{\delta \rightarrow 0} \int_{q_i}^{q_{i+1}} \frac{-\partial_r \mathcal{P}_{\psi,g}(x_\delta)}{\delta} \\ &= \int_{q_i}^{q_{i+1}} \lim_{\delta \rightarrow 0} \frac{-\partial_r \mathcal{P}_{\psi,g}(x_\delta)}{\delta}. \end{aligned}$$

The equality of left and right derivatives is checked using the fact that  $\lim_{\delta \rightarrow 0} x_{\delta,q_{i+1} \rightarrow r} = \lim_{\delta \rightarrow 0} x_{\delta,q_i \rightarrow r}$  and the  $x$ -continuity in Corollary 5.  $\square$

The simplicity of the differentiation scheme gives a formal approach for computing derivatives of all orders in  $x$  for RPC's functionals.

**4.3. Continuity and Monotonicity.** As a direct application of the differentiation formulas, we establish the  $L^1$ -continuity and the monotonicity of  $\mathcal{P}_{\psi,g}$ . This is a theorem due to Guerra [6] that we prove in a general setting.

The space  $\mathcal{M}$  has a natural partial ordering. We say that  $x \in \mathcal{M}$  *dominates*  $y \in \mathcal{M}$  if  $x(q) \geq y(q)$  for all  $q \in [0, 1]$ . The terminology refers to the stochastic dominance of the random variables associated to the distribution functions. A functional  $\Lambda : \mathcal{M} \rightarrow \mathbb{R}$  is said to be *monotone increasing* with respect to this partial ordering if for any  $x, y \in \mathcal{M}$  such that  $x$  dominates  $y$ ,  $\Lambda(x) \geq \Lambda(y)$ .

We start by stating a useful lemma whose proof is straightforward.

**Lemma 11.** *Let  $x, y \in \mathcal{M}_a$  and  $\Lambda : \mathcal{M}_a \rightarrow \mathbb{R}$ . Suppose  $\partial_{x_i} \Lambda$  exists for all  $x \in \mathcal{M}_a$ . Suppose also that the estimate  $0 \leq \partial_{x_i} \Lambda(x) \leq C_\Lambda (g(q_{i+1}) - g(q_i))$  holds for some  $C_\Lambda > 0$  and for a strictly increasing function  $g \in C^1([0, 1])$  such that  $g(0) = 0$ . Then*

$$(14) \quad \Lambda(x) - \Lambda(y) \leq C_\Lambda \int_0^1 \max\{x(q) - y(q), 0\} dg(q).$$

The theorem of Guerra appeared in [6] without proofs. A proof can be found in [14]. In the RPC formulation, it is a basic consequence of the bound on the  $x$ -derivative.

**Theorem 12** (Continuity and Mononicity). *Let  $\psi \in \mathcal{C}$  such that  $|\psi'| \leq C$  and  $g \in C^1([0, 1])$ , a strictly increasing function with  $g(0) = 0$ . Then the following holds:*

(1) *If  $x, y \in \mathcal{M}_a^{<1}$ ,*

$$|\mathcal{P}_{\psi,g}(x) - \mathcal{P}_{\psi,g}(y)| \leq \frac{C^2}{2} \|x - y\|_{L^1(g'(q)dq)}.$$

*In particular,  $\mathcal{P}_{\psi,g}$  has a continuous extension to the whole set  $\mathcal{M}$  in the  $L^1$ -topology (and so in the weak topology).*

(2)  *$\mathcal{P}_{\psi,g}$  is monotone increasing on  $\mathcal{M}$ .*

*Proof.* The bounds of Proposition 10 satisfy the assumptions of Lemma 11. The estimate of the first claim follows from equation (14). The  $L^1$ -continuity is clear from the estimate. The continuous extension is possible as  $\mathcal{M}_a^{<1}$  is dense subset of  $\mathcal{M}$  on the  $L^1$ -topology. The monotonicity is proven by Lemma 11 and the fact that if  $x, y \in \mathcal{M}_a$  and  $x(q) \geq y(q)$  for every  $q \in [0, 1]$ , then  $\max\{0, y(q) - x(q)\} = 0$ . The property is extended to  $\mathcal{M}$  by continuity.  $\square$

**4.4. The singular case  $x = 1$ .** When introducing the REM, we noticed that the case  $x = 1$  was singular as far as the intensity measure is concerned. However, it is necessary in the SK model theory to consider functionals on GREM's whose last level of splitting consist formally of REM(1). These GREM's correspond to elements of  $\mathcal{M}_a$  not in  $\mathcal{M}_a^{<1}$ . Using  $L^1$ -continuity, we are able to obtain an expression for  $\mathcal{P}_{\psi,g}$  and its differentiation formulas (9), (10) and (13) evaluated at these singular GREM's.

Let  $x \in \mathcal{M}_a$  with  $k$  atoms at  $\{q_l\}$  and  $q_k < 1$ . Then  $x \notin \mathcal{M}_a^{<1}$ . Consider  $x^\epsilon \in \mathcal{M}_a^{<1}$  with  $x^\epsilon(q_l) = x(q_l)$  for  $1 \leq l < k$ ,  $x^\epsilon(q_k) = 1 - \epsilon$  and  $x^\epsilon(q_{k+1}) = 1$ . Clearly,  $x^\epsilon \rightarrow x$  in  $L^1(g'(q)dq)$  as  $\epsilon$  goes to 0. Applying quasi-stationarity to a REM( $1 - \epsilon$ ) with shift  $e^{\psi(\kappa_\alpha(1))}$ , we pick up the factor  $\mathbb{E}_{x^\epsilon} [e^{(1-\epsilon)\psi(\kappa_\alpha(1))} | \mathcal{F}_{q_k}]^{1/1-\epsilon}$  which simply becomes in the limit :  $\mathbb{E}_{x^\epsilon} [e^{\psi(\kappa_\alpha(1))} | \mathcal{F}_{q_k}] =: e^{\psi_{q_k}(\kappa_{\alpha(k-1)}(q_k))}$ . One applies  $L^1$ -continuity and the dominated convergence theorem to  $\mathcal{P}_{\psi,g}(x^\epsilon)$  to get

$$\mathcal{P}_{\psi,g}(x) = \mathbb{E}_x \left[ \log \frac{\sum_{\alpha(k-1)} \xi_{\alpha(k-1)} e^{\psi_{q_k}(\kappa_{\alpha(k-1)}(q_k))}}{\sum_{\alpha(k-1)} \xi_{\alpha(k-1)}} \right].$$

The cases  $\psi(\eta) = \log \cosh(\beta\eta + h)$  and  $\psi(\eta) = \beta\eta$  are again special as

$$e^{\psi_{q_k}(\kappa_{\alpha(k-1)}(q_k))} := E_z \left[ e^{\psi(z\sqrt{g(1)-g(q_k)} + \kappa_{\alpha(k-1)}(q_k))} \right] = e^{\frac{\beta^2}{2}(g(1)-g(q))} e^{\psi(\kappa_{\alpha(k-1)}(q_k))}$$

and  $\psi$  is retrieved after integration. In these cases,

$$(15) \quad \mathcal{P}_{\psi,g}(x) = \frac{\beta^2}{2}(g(1) - g(q)) + \mathbb{E}_x \left[ \log \frac{\sum_{\alpha(k-1)} \xi_{\alpha(k-1)} e^{\psi(\kappa_{\alpha(k-1)}(q_k))}}{\sum_{\alpha(k-1)} \xi_{\alpha(k-1)}} \right].$$

In particular, if  $k = 1$ , i.e.  $x$  has a single atom sitting at  $q$ :

$$(16) \quad \mathcal{P}_{\psi,g}(x) = \frac{\beta^2}{2}(g(1) - g(q)) + \mathbb{E}_x [\psi(\kappa_{\alpha(0)}(q))] = \frac{\beta^2}{2}(g(1) - g(q)) + \int_{\mathbb{R}} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \psi(z\sqrt{g(q)}) dz$$

which, not surprisingly, resembles the high-temperature solution of the SK model.

The derivative formulas (9), (10) and (13) are retrieved by applying the derivative to the expression (15). The formulas hold by replacing  $\tilde{\kappa}_i(1)$  by  $\tilde{\kappa}_i(q_k)$ . As an example, we compute  $\partial_{q_k} \mathcal{P}_{\beta,h}(x)$ :

$$(17) \quad \begin{aligned} \partial_{q_k} \mathcal{P}_{\beta,h}(x) &= -g'(q_k) \frac{\beta^2}{2} + g'(q_k) \frac{\beta^2}{2} \mathbb{E}_x [(\psi'' + \psi'^2)(\tilde{\kappa}_1(q_k))] \\ &\quad - g'(q_k) \frac{\beta^2}{2} \mathbb{E}_x [\psi'(\tilde{\kappa}_1(q_k)) \psi'(\tilde{\kappa}_2(q_k)) \chi_{q_{12}=q_k}] \\ &= -g'(q_k) \frac{\beta^2}{2} \mathbb{E}_x [(\psi')^2(\tilde{\kappa}_1(q_k)) \chi_{q_{12}=q_k}] \end{aligned}$$

where we used the differentiation formula (25) in the first equality and the identity  $\psi'' + \psi'^2 = 1$  in the second.

## 5. THE $AS^2$ -VARIATIONAL PRINCIPLE AND THE PARISI FORMULA

The  $AS^2$  variational principle expresses the pressure of the SK model in the thermodynamic limit  $P_{SK}(\beta, h)$  as an optimization problem over random overlap structures [2, 3]:

$$(18) \quad P_{SK}(\beta, h) = \lim_{M \rightarrow \infty} \inf_{\{\mu \text{ ROST}\}} G_M(\beta, h, \mu)$$

where  $\mu$  is the probability measure of a ROST  $(\xi, Q)$ . The functional  $G_M$  is given by

$$G_M(\beta, h, \mu) := \frac{1}{M} \mathbb{E}_\mu \left[ \log \frac{\sum_\alpha \xi_\alpha \prod_{i=1}^M 2 \cosh(\beta \eta_\alpha^i + h)}{\sum_\alpha \xi_\alpha \prod_{i=1}^M e^{\beta \kappa_\alpha^i}} \right]$$

where  $\eta^i$  and  $\kappa^i$ ,  $i = 1, \dots, M$ , are independent copies of the cavity field and of a gaussian field with covariance function  $q^2/2$ .

As we noticed previously, the RPC's form a particularly interesting class of ROST's due to the ultrametric structure of the overlap matrix and the quasi-stationarity property. In this section, we study the  $AS^2$  variational principle (18) restricted to the class of RPC's. First, we show that the restricted variational problem reduces to the Parisi formula. Then, we rederive the Almeida-Thouless line which yields sufficient condition for the minimizer of the variational problem not to be a single atom. For this, we follow the idea of Toninelli but we explicitly use the RPC structure underlying the functionals.

**5.1. The Parisi Formula.** The Parisi formula for the pressure of the SK model was proven by Talagrand in [13].

**Theorem 13** (The Parisi Formula).

$$P_{SK}(\beta, h) = \inf_{x \in \mathcal{M}_a} \left\{ \log 2 + \mathcal{P}_{\beta, h}^{par}(x) - \frac{\beta^2}{2} \int_0^1 qx(q) dq \right\}.$$

It is remarked in [2, 3] that the Parisi formula is exactly the  $AS^2$  variational problem restricted to the class of RPC's. To establish this connection, we start by noting that the limit  $M \rightarrow \infty$  in equation (18) is no longer needed when we deal with RPC's.

**Proposition 14** (Variational Principle over the class of GREM's). *Let  $\mu$  be a GREM parametrized by  $x \in \mathcal{M}_a^{<1}$ . Then, for any  $M \in \mathbb{N}$ ,  $G_M(\beta, h, \mu) = G_1(\beta, h, \mu)$  and*

$$G_{\beta, h}(x) := G_1(\beta, h, \mu) = \mathbb{E}_x \left[ \log \frac{\sum_\alpha \xi_\alpha 2 \cosh(\beta \eta_\alpha(1) + h)}{\sum_\alpha \xi_\alpha} \right] - \mathbb{E}_x \left[ \log \frac{\sum_\alpha \xi_\alpha e^{\beta \kappa_\alpha(1)}}{\sum_\alpha \xi_\alpha} \right].$$

The variational problem of equation (18) restricted to GREM's reduces to

$$\lim_{M \rightarrow \infty} \inf_{\mu \in \text{GREM}} G_M(\beta, h, \mu) = \inf_{x \in \mathcal{M}_a} G_{\beta, h}(x).$$

*Proof.* It suffices to note that each of the  $M$  independent copy of the fields contributes the same factor to the pressure. This is done using equation (7).  $\square$

Note that  $G_{\beta, h}(x)$  is the difference of two RPC functionals with  $\psi(\eta) = \log \cosh(\beta \eta + h)$ ,  $g(q) = q$  and  $\psi(\kappa) = \beta \kappa$ ,  $g(q) = q^2/2$  respectively. The Parisi formula is retrieved from the  $AS^2$  variational principle on RPC's by using Theorem 6 and Proposition 7 for these two functionals:

**Proposition 15** (The Parisi Formula with RPC's).

$$(19) \quad P_{SK}(\beta, h) = \inf_{\text{GREM}(x)} G_{\beta, h}(x).$$

As  $G_{\beta, h}$  is the difference of two  $L^1$ -continuous functionals by Theorem 12, it is itself  $L^1$ -continuous and can be extended continuously to  $\mathcal{M}$ . Therefore we can rewrite equation (19) as

$$(20) \quad P_{SK}(\beta, h) = \min_{x \in \mathcal{M}} G_{\beta, h}(x)$$



because  $\mathcal{M}$  is compact in the  $L^1$ -topology (recall that the  $L^1$ -topology is equivalent to the weak topology. See Appendix A). The question of the uniqueness of the minimizer of the Parisi formula was raised in [10, 14] but remains open.

**5.2. An example of calculation: The Almeida-Thouless Line.** It is now well known that the minimizer of the variational principle (20) for the SK model is an atomic measure with a single atom when  $\beta$ , as a function of  $h$ , is small enough [1, 9]. This is referred to as the high-temperature solution of the SK model. Moreover, it was proven by Toninelli that this solution cannot hold beyond the so-called Almeida-Thouless line [15]. In this section, we rederive this sufficient condition following Toninelli in spirit, but relying heavily on the RPC structure. In doing so, we hope to illustrate the convenient features of the RPC formalism<sup>1</sup>.

First, we remark that the stationarity conditions of the optimization problem (20)  $\partial_{q_i} G_{\beta,h}(x) = 0$  for  $i = 0, 1, \dots, k$  yield self-consistency equations:

$$\mathbb{E}_x [\tanh(\beta\tilde{\eta}_1(1) + h) \tanh(\beta\tilde{\eta}_2(1) + h) | q_{12} = q_i] = q_i.$$

This is a consequence of the definition of  $G_{\beta,h}$  and Proposition 8. In the case of a single atom, we get the self-consistency equation for the high-temperature solution from equation (17)

$$E_z [\tanh^2(\beta z \sqrt{q} + h)] = q.$$

We denote a solution to the self-consistency equation  $\bar{q} = \bar{q}(\beta, h)$ . This solution is unique in the case  $\beta < 1$  and  $h = 0$  and in the case  $h \neq 0$  (see e.g. [9]).

Fix  $\beta > 0$  and  $h \in \mathbb{R}$ . The high-temperature solution corresponds to the infimum of  $G_{\beta,h}$  over the subset of  $\mathcal{M}_a$  consisting of measures with a single atom. Note that this set is compact so the infimum is attained. Let  $x^* \equiv x^*(\beta, h)$  be the minimizer. The only atom of  $x^*$  must be located at  $\bar{q}$ , the solution to the self-consistency equation. The idea for deriving the Almeida-Thouless condition is to show that, if  $\beta$  and  $h$  are such that

$$\beta^2 \int_{\mathbb{R}} \frac{e^{-z^2/2}}{\sqrt{2\pi}} \cosh^{-4}(\beta z \sqrt{q} + h) > 1,$$

then there exists an element of  $\mathcal{M}_a$  with two atoms such that  $G_{\beta,h}$  evaluated at that element is smaller than  $G_{\beta,h}(x^*)$ . This implies that the high-temperature solution cannot hold for the optimization problem (20) in this region of the plane  $(\beta, h)$ .

Let us construct such an element. Pick  $0 \leq m \leq 1$  and  $\bar{q} \leq r \leq 1$ . Let  $x_{m,r}^*$  be the atomic measures with atoms at  $\bar{q}$  and  $r$  with  $x_{m,r}^*(\bar{q}) = m$  and  $x_{m,r}^*(r) = 1$ . Note that  $x_{m,\bar{q}}^* = x^*$ . As  $G_{\beta,h}(x_{m,r}^*)$  is continuously differentiable in  $m$  on  $0 < m < 1$  for  $r > \bar{q}$ , we have

$$(21) \quad G_{\beta,h}(x^*) = G_{\beta,h}(x_{m,r}^*) + \int_m^1 \partial_{m'} G_{\beta,h}(x_{m',r}^*) dm'.$$

From equation (21), we see that  $G_{\beta,h}(x_{m,r}^*) < G_{\beta,h}(x^*)$  if  $r$  and  $m$  are such that

$$(22) \quad \partial_{m'} G_{\beta,h}(x_{m',r}^*) > 0$$

<sup>1</sup>It was recently proven by Guerra that the high-temperature solution actually holds up to the Almeida-Thouless line. This had been rigorously established only in the case  $h = 0$  [8].

for  $m \leq m' < 1$ . To get this inequality, we follow [15] and expand  $\partial_m G_{\beta,h}(x_{m,r}^*)$  around  $\bar{q}$

$$(23) \quad \begin{aligned} \partial_m G_{\beta,h}(x_{m,r}^*) &= \partial_m G_{\beta,h}(x_{m,r}^*) \Big|_{r=\bar{q}} + (r - \bar{q}) \partial_r \partial_m G_{\beta,h}(x_{m,r}^*) \Big|_{r=\bar{q}} \\ &\quad + \frac{(r - \bar{q})^2}{2} \partial_r^2 \partial_m G_{\beta,h}(x_{m,r}^*) \Big|_{r=\bar{q}} + \mathcal{O}((r - \bar{q})^3). \end{aligned}$$

The remainder term is bounded (this can be checked using the gaussian differentiation formula to calculate the third derivative).

The first term of the expansion is 0 by Proposition 10 as the integral involved in the differentiation formula is from  $\bar{q}$  to  $r$ . The second term vanishes too as we retrieve the self-consistency equation for  $\bar{q}$  by differentiating in  $r$  the integral expression for  $\partial_m G_{\beta,h}(x_{m,r}^*)$

$$\partial_r \partial_m G_{\beta,h}(x_{m,r}^*) \Big|_{r=\bar{q}} = \mathbb{E}_{x_{m,r}^*} [\tanh(\beta \tilde{\eta}_1(r) + h) \tanh(\beta \tilde{\eta}_2(r) + h) | q_{12} = q_i] - r \Big|_{r=\bar{q}} = 0.$$

Therefore to prove (22), it suffices to find conditions for which  $\partial_m \partial_r^2 G_{\beta,h}(x_{m,r}^*) \Big|_{r=\bar{q}} > 0$ . Equation (11) is useful to compute  $\partial_r^2 G_{\beta,h}(x_{m,r}^*)$ . From Theorem 2, the matrices  $Q_3$  and  $Q_4$  with  $\mathbb{P}_{x_{m,r}^*}(Q_3) \neq 0$  and  $\mathbb{P}_{x_{m,r}^*}(Q_4) \neq 0$  are

$$\begin{aligned} \mathbb{P}_{x_{m,r}^*}(q_{12} = r) &= 1 - m \\ \mathbb{P}_{x_{m,r}^*}(q_{12} = r, q_{23} = r) &= \frac{(2 - m)(1 - m)}{2} \\ \mathbb{P}_{x_{m,r}^*}(q_{12} = r, q_{34} = r, q_{13} = r) &= \frac{(3 - m)(2 - m)(1 - m)}{6} \\ \mathbb{P}_{x_{m,r}^*}(q_{12} = r, q_{34} = r, q_{13} = \bar{q}) &= \frac{m(1 - m)^2}{6}. \end{aligned}$$

The reader can check that all missing overlaps  $q_{ij}$  of the matrices  $Q_3$  and  $Q_4$  in the above events are determined by ultrametricity. For example, if  $q_{12} = r, q_{23} = r$  then  $q_{13} = r$ . From the probabilities above, one can see that, when applying the derivative  $\partial_m$  directly to  $\partial_r^2 G_{\beta,h}(x_{m,r}^*)$  and taking the limit  $m \rightarrow 1^-$ , only the terms coming from the derivative of the factor  $1 - m$  do not vanish. Thus, one gets the remaining terms

$$(24) \quad \begin{aligned} \lim_{m \rightarrow 1^-} \left( \partial_r^2 \partial_m G_{\beta,h}(x_{m,r}^*) \Big|_{r=\bar{q}} \right) &= -\frac{\beta^2}{2} (1 - \beta^2 \mathbb{E}_{x^*} [1 - 2 \tanh^2(\beta(\tilde{\eta}(\bar{q}) + h)) + \tanh^4(\beta(\tilde{\eta}(\bar{q}) + h))]) \\ &= -\frac{\beta^2}{2} (1 - \beta^2 \mathbb{E}_{x^*} [\cosh^{-4}(\beta(\tilde{\eta}(\bar{q}) + h))]) \end{aligned}$$

where we have used the fact that  $x_{m,r}^* \Big|_{r=\bar{q}} = x^*$ . Hence by the Taylor's expansion (23) and the equation (24), if  $\beta$  and  $h$  are such that

$$\beta^2 \mathbb{E}_{x^*} [\cosh^{-4}(\beta(\tilde{\eta}(\bar{q}) + h))] = \beta^2 \int_{\mathbb{R}} \frac{e^{-z^2/2}}{\sqrt{2\pi}} \cosh^{-4}(\beta z \sqrt{\bar{q}} + h) > 1,$$

we can pick  $m$  close enough to 1 and  $r$  close enough to  $\bar{q}$  so that the inequality (22) holds. This yields the desired sufficient condition for the high-temperature solution not to hold.

APPENDIX A. TOPOLOGY ON  $\mathcal{M}$ 

The weak, vague and weak-\* convergences correspond on the space of probability measures  $\mathcal{M}$  as the measures are on a compact of  $\mathbb{R}$ . Moreover, any sequence is tight. Therefore, the weak topology is determined by the weak convergence.

In studying the functionals on  $\mathcal{M}$ , we are led to consider the topology on  $\mathcal{M}$  induced by the  $L^1([0, 1], g'(q)dq)$ -norm on the distribution functions of the elements of  $\mathcal{M}$  where  $g$  is a strictly increasing function in  $C^1([0, 1])$  with  $g(0) = 0$ . It turns out that all these norms induce topologies on  $\mathcal{M}$  that are equivalent to the weak topology.

First, we claim that the  $L^1(g'(q)dq)$ -topology is equivalent to the  $L^1(dq)$ -topology on the space  $\mathcal{M}$ . Clearly,  $\|\cdot\|_{L^1(g'(q)dq)} \leq \max_q g'(q) \|\cdot\|_{L^1(dq)}$ . On the other hand, the following estimate holds for any  $\delta > 0$ :

$$\|\cdot\|_{L^1(dq)} \leq \frac{1}{\delta} \|\cdot\|_{L^1(g'(q)dq)} + \text{Leb}\{q \in [0, 1] : 0 \leq g'(q) < \delta\}$$

where  $\text{Leb}$  stands for the Lebesgue measure. We use the fact that the distribution functions are bounded above by 1 and below by 0 to get the second term. As  $g$  is strictly increasing, we have that, for any  $\epsilon > 0$ , there exists  $\delta(\epsilon)$  such that  $\text{Leb}\{q \in [0, 1] : 0 \leq g'(q) < \delta\} < \epsilon$ . Let  $x_\gamma$  be a net of distribution functions on  $[0, 1]$  that converges in the  $L^1(g'(q)dq)$ -norm. To see that  $x_\gamma$  also converges in the  $L^1(dq)$ -norm, it suffices to see that for  $\delta$  arbitrary small but fixed, one can also make the first term of the r.h.s. of the above estimate arbitrary small using the convergence in the  $L^1(g'(q)dq)$ -norm.

The equivalence with the weak topology is a direct consequence of the fact that the  $L^1(dq)$ -norm metrizes the weak topology on  $\mathcal{M}$  (see [16]).

## APPENDIX B. GAUSSIAN DIFFERENTIATION FORMULAS

The differentiation of expectations of gaussian variables whose covariance depends on a parameter is facilitated by the following result which can be seen as an extension of the Wick's formula or simply gaussian integration by parts.

**Proposition 16** ([2, 3]). *Consider a gaussian vector  $\kappa = (\kappa_i, i \in \mathbb{N})$  for which the covariance matrix  $\{c_{ij}(t)\}$  depends on a parameter  $t \in [0, 1]$ . Assume  $c_{ij} \in C^1([0, 1])$ . We write  $\mathbb{E}_t$  for the expectation over  $\kappa$ . Let  $\phi : \mathbb{R}^N \rightarrow \mathbb{R}$  in  $C^2(\mathbb{R}^N)$  whose derivatives multiplied by  $e^{-\epsilon|x|^2}$  are bounded functions for any  $\epsilon > 0$ . Then*

$$\frac{d}{dt} \mathbb{E}_t[\phi(\kappa)] = \frac{1}{2} \sum_{i,j} c'_{ij}(t) \mathbb{E}_t [\partial_{\kappa_i} \partial_{\kappa_j} \phi(\kappa)].$$

The proof of the proposition is easy to carry for polynomials (this case is the usual Wick's formula). For the general case, we refer to [3] for a proof using the Fourier transform.

In the case  $\phi(\kappa) = \log \sum_\alpha \xi_\alpha e^{\psi(\kappa_\alpha)}$  for a set of weight  $\{\xi_\alpha\}$ , the formula becomes

$$(25) \quad \begin{aligned} \frac{d}{dt} \mathbb{E}_t[\phi(\kappa)] &= \frac{1}{2} \sum_\alpha c'_{\alpha,\alpha}(t) \mathbb{E}_t [\psi''(\kappa_\alpha) + \psi'^2(\kappa_\alpha)] \frac{\xi_\alpha e^{\psi(\kappa_\alpha)}}{\sum_\alpha \xi_\alpha e^{\psi(\kappa_\alpha)}} \\ &\quad - \frac{1}{2} \sum_{\alpha,\alpha'} c'_{\alpha,\alpha'}(t) \mathbb{E}_t [\psi'(\kappa_\alpha) \psi'(\kappa_{\alpha'})] \frac{\xi_\alpha \xi_{\alpha'} e^{\psi(\kappa_\alpha)} e^{\psi(\kappa_{\alpha'})}}{\sum_{\alpha,\alpha'} \xi_\alpha \xi_{\alpha'} e^{\psi(\kappa_\alpha)} e^{\psi(\kappa_{\alpha'})}}. \end{aligned}$$

It is possible to get higher-order derivatives by just applying the above gaussian differentiation formula successively. As an example, if the covariance  $c_{ij}$  depends linearly on

two parameters  $s$  and  $t$ ,  $c_{ij} = c_{ij}(s, t)$ , then applying the formula (16) twice yields

$$(26) \quad \partial_s \partial_t \mathbb{E}_{s,t}[\phi(\kappa)] = \frac{1}{4} \sum_{i,j,k,l} \partial_s c_{ij} \partial_t c_{ij} \mathbb{E}_{s,t} [\partial_{\kappa_i} \partial_{\kappa_j} \partial_{\kappa_k} \partial_{\kappa_l} \phi(\kappa)] .$$

#### REFERENCES

- [1] Aizenman M., Lebowitz J., Ruelle D., *Some rigorous results on the Sherrington-Kirkpatrick spin glass model*, Comm. Math. Phys. **112** (1987) pp. 3-20;
- [2] Aizenman M., Sims R., Starr S., *An Extended Variational Principle for the SK Spin-Glass Model*, Phys. Rev. B **68** (2003), pp. 214403;
- [3] Aizenman M., Sims R., Starr S., *Mean Field Spin Glass Models from the Cavity-ROSt Perspective*, preprint 2006, arXiv:math-ph/0607060;
- [4] Bolthausen E., Sznitman A.-S., *On Ruelle's Probability Cascades and an Abstract Cavity Method*, Comm. Math. Phys. **197** (1998) pp. 247-276;
- [5] Derrida B., *Random-energy model: Limit of a family of disordered models*, Phys. Rev. Lett., **45** (1981) pp. 79-82; Derrida B., *Random-energy model: An exactly solvable model of disordered systems*, Phys. Rev. B, **24** (1981) pp. 2613-2626; Derrida B. *A generalization of the random energy model which includes correlations between energies*, J. Phys. Lett. **46** (1985), pp. L401-L407;
- [6] Guerra F., *Broken Replica Symmetry Bounds in the Mean Field Spin Glass Model*, Comm. Math. Phys. **233** (2003) pp.1-12;
- [7] Guerra F., *About the cavity fields in mean field spin glass models*, arXiv.org:cond-mat/0307673
- [8] Guerra F., *The replica symmetric region in the Sherrington-Kirkpatrick mean field spin glass model. The Almeida-Thouless line*, arXiv.org: cond-mat/0604674;
- [9] Guerra F., Toninelli F. L., *Quadratic replica coupling in the Sherrington-Kirkpatrick mean field spin glass model*, J. Math. Phys. **43** (2002) pp. 3704;
- [10] Panchenko D., *A question about Parisi functional*, arXiv.org:/math.PR/0412463;
- [11] Ruelle D., *A Mathematical Reformulation of Derrida's REM and GREM*, Comm. Math. Phys. **108** (1987) pp. 225-239;
- [12] Ruzmaikina A., Aizenman M., *Characterization of invariant measures at the leading edge for competing particle systems*, Ann. Probab. **33** (2005), pp.82113;
- [13] Talagrand M., *The Parisi Formula*, Ann. Math. **163** (2006), pp. 221-263;
- [14] Talagrand M., *Parisi measures*, J. Func. Anal., to appear;
- [15] Toninelli F., *About the Almeida-Thouless transition line in the Sherrington-Kirkpatrick mean field spin glass model*, arXiv.org: cond-mat/0207296;
- [16] Villani C., *Topics in Optimal Transportation*, AMS, Providence (2003), 370 pp.

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON NJ 08544